

## **Small Area Estimate Program (SAEP) Methodology Overview**

### **Philosophy**

As a data driven process, the Small Area Estimates Program (SAEP) uses the best available data for each estimate area at the time of the data run. Therefore, estimates will change from run to run depending on when new data are added or updated for a given area. Changes apply to the current year's estimate as well as to historical estimates. Run to run changes may be large and are dependent on changes in the control data.

The estimate process is broken into three distinct parts: group quarter estimation, housing unit estimation, and household population estimation. The estimates themselves are shown with three decimal places of precision which smoothes the raking process and reminds the user that this is an estimate, not a hard count. The process is run year by year, with each new estimate building off the prior year's estimate.

### **Base Data**

Base data for SAEP comes from the 2000 census block data and the Office of Financial Management's (OFM) existing city/county population estimation process. Data from these sources is used as a foundation on which population and housing change is based.

OFM's official April 1 estimates of population and housing are used as controls throughout the estimation process. CenEst, OFM's census estimation database and system, is an in-house system used to estimate our annual county and city population as well as to track annexations and de-annexations. CenEst is the source of the control numbers of population, housing and group quarter populations used in the small area estimate process.

Geographic boundaries of cities continuously change through annexation and incorporations. Annexations are difficult to track as many are small in terms of area and size of population. In Washington, any city annexing or de-annexing an area is required to conduct a door-to-door census of population and housing in the annexation area. Annexation censuses are reviewed and certified as part of OFM's city/county population estimation process. Changes to city/county population are included in the control numbers.

Changes in jurisdiction boundaries resulting from annexation need to be accounted for geographically. It is for that reason we developed what we call "SAEP tabulation areas" as part of the small area estimates program. SAEP tabulation areas are approximations of current city limits and unincorporated balance of county areas defined by groups of census blocks. Like vintage 2000 census boundaries, tabulation areas are made up of complete census blocks; each block is entirely inside or outside city limits. SAEP tabulation area limits are determined by spatially joining annual city limit boundary files from the Washington State Department of Transportation (WSDOT) to census 2000 census blocks. After the joins are complete the percent area that each city contributes to each census block is determined according to the WSDOT city boundaries. If the newly annexed boundaries cover the majority of the area of a given census block, then that block is considered to be part of the city and not part of the unincorporated county for that estimate year. SAEP city limits and the balance of the county become the control areas utilized throughout the estimate process. New incorporations are assigned geographic areas in a similar manner.

## **Group Quarters Population Allocation**

Group Quarters (GQ) population is tracked from a variety of data sources. Cities report April 1 GQ populations by facility annually as part of OFM's annual population and housing survey. State agencies, like the Department of Corrections and Department of Social and Health Services, report populations for their facilities. The data for individual facilities are entered into the CenEst database and tracked over time.

Change by facility is calculated by subtracting the base year's (usually 2000) administrative population count from the current year's administrative population count.

These population changes are then applied to the appropriate census block. For SAEP, the GQ population of any individual block is not allowed to go below zero. Any change that cannot be allocated to individual census blocks is allocated on a proportional basis to all of the census blocks that contribute to the jurisdiction GQ total population. Note: There are a group of facilities for which we receive a total population by facility, but the actual census counts are on multiple blocks (colleges, some military bases etc.). Changes in population for these facilities are distributed to census blocks based on the proportions of populations counted in the 2000 census. In some cases the jurisdiction as a whole shows a loss in GQ population. When this happens the loss in population comes out of the household population.

## **Housing Unit Allocation**

The allocation of housing units to small areas is a data driven process. Housing change data are extracted from CenEst. A master data set that has the number of new units and demolitions by type (single family, multi family, mobile homes and specials) is created. Housing counts from CenEst serve as the control numbers for census blocks in the SAEP.

For each jurisdiction, each years completed/demolished housing units are distributed inside of the appropriate SAEP tabulation area boundaries based on the distribution of units from up to three sources: geocoded building permits, postal delivery estimates, or the census 2000 housing counts, depending on which data source(s) is available and the most accurate. Through each point in the process units are checked to see if the available data distribution is appropriate for use.

Geocoded building permits are considered the most accurate geographic indicator of housing growth/change. In most cases permits can be located with a high level of geographic accuracy and geocoded permits are given the first preference in housing allocation determination. If the control number of permits from CenEst is greater than 60 percent of the number of geocoded permits by type, then the control units are distributed to census blocks based on this distribution of geocoded permits. If the control number of building permits from CenEst is less than 60 percent of the total number of geocoded building permits then a pool of un-allocated geocoded permits plus the current years geocoded permits are checked to see if that pool of geocoded permits matches up with the cities control numbers of completions. If the base year plus carry over pool of geocoded permits is 60 percent of the CenEst control permits, then the data are allocated based on this distribution. If there is no reasonable match between the number of geocoded building permits and the CenEst control counts, then the data are run through the postal data allocation process.

Postal delivery data are the second level source of housing change and are considered less accurate than geocoded data because delivery statistics are available by the postal carrier route. These postal carrier routes can change drastically over time and can be large in rural areas. Routes do not cover the entire state nor do routes exist in areas where there is no mail delivery, i.e. areas with very low populations or where the primary source of mail delivery is via P.O. Boxes. To mitigate the effects of these boundary problems, the carrier route data are allocated to a stable geographic entity, the census block. Active and possible deliveries are proportionally allocated to census blocks based on the estimated number of housing units within each postal carrier route on an annual basis. As a check to see how well the postal data matches the federal census data, each jurisdiction's April 2000 possible delivery count is compared to the federal census count of housing units. This comparison is done in a statistic we refer to as the HouseHold Match rate or HHMatch rate. If a city's HHMatch rate is 60 percent or higher, the postal allocation data are used to distribute housing units within the jurisdiction. If the HHMatch rate is less than 60 percent, the census 2000 data are used to distribute the control figures.

Remaining housing data are allocated to blocks within a given jurisdiction based on the 2000 census distribution of housing. This generally occurs in small towns in rural areas. In many cases the city limit is significantly smaller than the surrounding census tract and is therefore a reasonable boundary to limit the placement of units. The remaining units are simply proportionally allocated based on the 2000 census distribution of occupied housing units. This process still leaves some jurisdictions over or under allocated. Jurisdictions with counts that do not match the control figures have their balances raked until the unit counts match. The result of this process is a census block estimate of housing by year. It is important to note that this allocation process is done by structure type and is done separately for new constructions and demolitions.

Very often our geocoded building permit data can be used as a reasonable means to distribute new housing, but there is usually a lack of geocoded demolitions. Subsequently, many jurisdictions permits are allocated out based on geocoded permits, but demolitions are allocated based on the postal change or less frequently by federal census data.

## **Population Allocation**

After the housing units are allocated to blocks, a master occupancy and average Population Per Household (PPH) table is extracted from CenEst. For each year, the change in occupancy and PPH is calculated on a jurisdiction-wide basis. Occupancy rates and PPH ratios are also calculated at the census block level based on the 2000 census data. The jurisdiction-wide change in occupancy rates and PPH ratios is applied to all census blocks within their appropriate SAEP tabulation area. If the mathematical result of adding the change in occupancy rate to the base census rate results in an occupancy rate above one, the occupancy rate is set to one. If the application of change would result in an occupancy rate of less than zero, then the block rate is left at the base census rate.

For existing housing (i.e. units which are not new this estimate year) the number of occupied housing units is estimated by applying the new adjusted block occupancy rates to the existing block housing counts. For new housing units, occupied units are estimated by applying the jurisdiction average occupancy rate for the appropriate housing type. This is done separately for existing housing and for newly added units because new units sometimes fall in blocks where there was little or no existing housing. The adjusted block rates do not apply in these cases.

Output from the housing unit allocation process is used to estimate household population. This is done separately for existing housing and for newly added units similar to the way the occupancy rates were applied to the estimate housing stock. A population base estimate is made for existing housing by applying estimated block PPH ratios to estimated occupied units. Population from new housing units is estimated using the average jurisdiction PPH ratio by structure type. In our city estimates we often adjust the occupancy rates for jurisdictions with large numbers of new units to reflect the lower occupancy rates associated with large new developments. A similar adjustment is made in the SAEP as well. High growth census blocks, those with more than 25 new units per year, have their occupancy rates decreased for new units by ten percent the first year to help compensate for the lower occupancy rates associated with high growth developments.

The total population for the jurisdiction is then calculated using the years base population, the population estimated from new housing units and the group quarter population estimate. The difference from the control population is calculated and is then raked over the jurisdiction based on the occupied housing stock (base and new). When the raking process is complete, occupancy rates and PPH ratios are re-calculated at the block level based on the raked data. These values become the base data that feed the next year's population allocation process. Then the population allocation process starts over for the next year.